



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

| | | |
|---|----|---|
| (51) International Patent Classification ⁷ : | A1 | (11) International Publication Number: WO 00/67162 |
| G06F 17/30 | | (43) International Publication Date: 9 November 2000 (09.11.00) |

(21) International Application Number: PCT/US00/12386

(22) International Filing Date: 5 May 2000 (05.05.00)

(30) Priority Data:
60/132,673 5 May 1999 (05.05.99) US

(71) Applicant (for all designated States except US): WEST PUBLISHING COMPANY [US/US]; d/b/a West Group, 610 Opberman Drive, Eagan, MN 55123 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): YANG-STEVENS, Bokyung [KR/US]; 13955 Granada Court, Apple Valley, MN 55124 (US). SWOPE, M., Charles [US/US]; 79 MacKubin Street, St. Paul, MN 55102 (US). LOCKE, Jeffrey [US/US]; 310 East Logan Avenue, West St. Paul, MN 55118 (US). MOULINIER, Isabelle [FR/US]; Apartment 1208, 3480 Golfview Drive, Eagan, MN 55123 (US).

(74) Agent: VIKSNINS, Ann, S.; Schwegman, Lundberg, Woessner & Kluth, P.O. Box 2938, Minneapolis, MN 55402 (US).

(81) Designated States: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

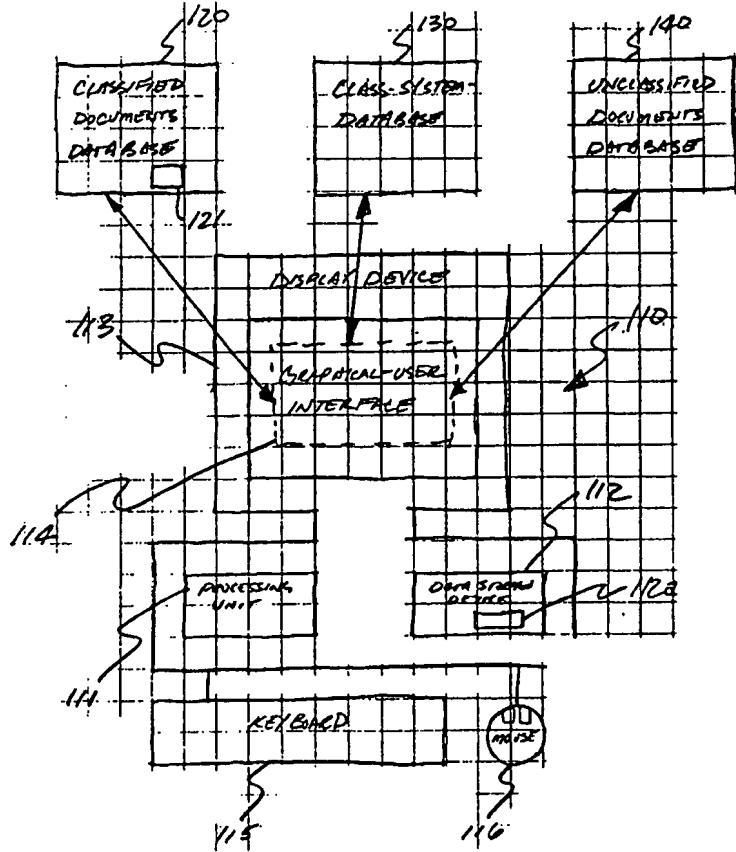
Published

With international search report.

(54) Title: DOCUMENT-CLASSIFICATION SYSTEM, METHOD AND SOFTWARE

(57) Abstract

Every year, professional classifiers at West Group manually classify over 350,000 headnotes, or abstracts of judicial opinions across approximately 82,000 separate classes of the Key Number System (130). Although most headnotes are classified from the memory of classifiers, a significant number are difficult and thus costly to classify (130) manually. Accordingly, the inventors devised systems (120), methods, and software that facilitate manual classification (120) of headnotes and documents generally hard-to-classify and particularly headnotes. One exemplary system provides a graphical user interface (114) that concurrently displays an unclassified headnote (140), a ranked list of one or more candidate classes, a candidate class in combination with adjacent classes of the classification system (100), and at least one classified headnote associated with one of the candidate classes.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | | | | | | | |
|----|--------------------------|----|---------------------------------------|----|---|----|--------------------------|
| AL | Albania | ES | Spain | LS | Lesotho | SI | Slovenia |
| AM | Armenia | FI | Finland | LT | Lithuania | SK | Slovakia |
| AT | Austria | FR | France | LU | Luxembourg | SN | Senegal |
| AU | Australia | GA | Gabon | LV | Latvia | SZ | Swaziland |
| AZ | Azerbaijan | GB | United Kingdom | MC | Monaco | TD | Chad |
| BA | Bosnia and Herzegovina | GE | Georgia | MD | Republic of Moldova | TG | Togo |
| BB | Barbados | GH | Ghana | MG | Madagascar | TJ | Tajikistan |
| BE | Belgium | GN | Guinea | MK | The former Yugoslav Republic of Macedonia | TM | Turkmenistan |
| BF | Burkina Faso | GR | Greece | | | TR | Turkey |
| BG | Bulgaria | HU | Hungary | ML | Mali | TT | Trinidad and Tobago |
| BJ | Benin | IE | Ireland | MN | Mongolia | UA | Ukraine |
| BR | Brazil | IL | Israel | MR | Mauritania | UG | Uganda |
| BY | Belarus | IS | Iceland | MW | Malawi | US | United States of America |
| CA | Canada | IT | Italy | MX | Mexico | UZ | Uzbekistan |
| CF | Central African Republic | JP | Japan | NE | Niger | VN | Viet Nam |
| CG | Congo | KE | Kenya | NL | Netherlands | YU | Yugoslavia |
| CH | Switzerland | KG | Kyrgyzstan | NO | Norway | ZW | Zimbabwe |
| CI | Côte d'Ivoire | KP | Democratic People's Republic of Korea | NZ | New Zealand | | |
| CM | Cameroon | | | PL | Poland | | |
| CN | China | KR | Republic of Korea | PT | Portugal | | |
| CU | Cuba | KZ | Kazakhstan | RO | Romania | | |
| CZ | Czech Republic | LC | Saint Lucia | RU | Russian Federation | | |
| DE | Germany | LI | Liechtenstein | SD | Sudan | | |
| DK | Denmark | LK | Sri Lanka | SE | Sweden | | |
| EE | Estonia | LR | Liberia | SG | Singapore | | |

DOCUMENT-CLASSIFICATION SYSTEM, METHOD AND SOFTWARE

5

Cross-Reference to Related Applications

This application is a continuation of U.S. provisional patent application 60/132673 which was filed May 5, 1999 and which is incorporated herein by reference.

10

Copyright Notice and Permission

A portion of this patent document contains material subject to copyright protection. The copyright owner has no objection to the facsimile reproduction by anyone of the patent document or the patent disclosure, as it appears in the Patent and Trademark Office patent files or records, but otherwise reserves all 15 copyright whatsoever. The following notice applies to this document:

Copyright © 1999, West Group

Technical Field

The present invention concerns document classification systems and methods for legal documents, such as judicial decisions.

20

Background

The American legal system, as well as some other legal systems around the world, relies heavily on written judicial opinions ---the written pronouncements of judges--- to articulate or interpret the laws governing resolution of disputes. Each judicial opinion is not only important to resolving a 25 particular dispute, but also to resolving all similar disputes in the future. This importance reflects the principle of American law that the judges within a given jurisdiction should decide disputes with similar factual circumstances in similar ways. Because of this principle, judges and lawyers within the American legal system are continually searching an ever-expanding body of past decisions, or 30 case law, for the decisions that are most relevant to resolution of particular disputes.

To facilitate this effort, companies, such as West Group (formerly West Publishing Company) of St. Paul, Minnesota, not only collect and publish the judicial opinions of jurisdictions from almost every federal and state jurisdiction 35 in the United States, but also classify the opinions based on the principles or

points of law they contain. West Group, for example, classifies judicial opinions using its proprietary Key Number™ System. (Key Number is a trademark of West Group.) This system has been a seminal tool for finding relevant judicial opinions since the turn of the century.

5 The Key Number System is a hierarchical system of over 400 major legal topics, with the topics divided into subtopics, the subtopics into sub-subtopics, and so on. Each topic or sub-topic has a unique alpha-numeric code, known as its Key Number classification. Table 1 shows an example of a portion of the Key Number System for classifying points of divorce law:

| 10 | Key Number Classification | Topic Description |
|----|---------------------------|---|
| | 134 | Divorce |
| | 134V | Alimony, Allowances, and Property Disposition |
| | 134k230 | Permanent Alimony |
| | 134k235k | Discretion of Court |

15 Table 1. Key Number hierarchy and corresponding Topic Descriptions

At present, there are approximately 82,000 Key Number classes or categories, each one delineating a particular legal concept.

20 Maintaining the Key Number System is an enormous on-going effort, requiring hundreds of professional editors to keep up with the thousands of judicial decisions issued throughout the United States ever year. Professional attorney-editors read each opinion and annotate it with individual abstracts, or headnotes, for each point of law it includes. The resulting annotated opinions 25 are then passed in electronic form to classification editors, or classifiers, who read each headnote and manually assign it to one or more classes in the Key Number System. For example, a classifier facing the headnote: "Abuse of discretion in award of maintenance occurs only where no reasonable person would take view adopted by trial court assigned." would most likely assign it to 30 Key Number class 134k235, which as indicated in Table 1, corresponds to the Divorce subtopic "discretion of court".

Every year, West Group classifiers manually classify over 350,000 headnotes across the approximately 82,000 separate classes of the Key Number classification system. Over time, many of the classifiers memorize significant

portions of the Key Number System, enabling them to quickly assign Key Number classes to most headnotes they encounter. However, many headnotes are difficult to classify. For these, the classifier often invokes the WestLaw™ online legal search service, which allows the user to manually define queries 5 against a database of classified headnotes. (WestLaw is a trademark of West Group.)

For instance, if presented with the exemplary "abuse of discretion" headnote, an editor might define and run a query including the terms "abuse," "discretion," "maintenance," and "divorce." The search service would return a 10 set of annotated judicial opinions compliant with the query and the classifier would in turn sift through the headnotes in each judicial opinion, looking for those most similar to the headnote targeted for classification. If one or more of the headnotes satisfies the editor's threshold for similarity, the classifier manually assigns the Key Number classes associated with these headnotes to the 15 target headnote. The classifier, through invocation of a separate application, may also view an electronic document listing a portion of the Key Number System to help identify related classes that may not be included in the search results.

The present inventors recognized that this process of classification suffers 20 from at least two problems. First, even with use of online searching, the process is quite cumbersome and inefficient. For example, editors are forced to switch from viewing a headnote in one application, to a separate online search application to manually enter queries and view search results, to yet another application to consult a classification system list before finally finishing 25 classification of some hard-to-classify headnotes. Secondly, this conventional process of classification lacks an efficient method of correcting misclassified headnotes. To correct misclassified headnotes, a classifier makes a written request to a database administrator with rights to a master headnote database.

Accordingly, there is a need for systems, methods, and software that not 30 only streamline manual classification processes, but also promote consistency and accuracy of resulting classifications.

Summary

To address this and other needs, the inventors devised systems, methods, and software that facilitate the manual classification of documents, particularly judicial opinions according to a legal classification system, such as West

- 5 Group's Key Number System. One exemplary system includes a personal computer or work station coupled to a memory storing classified judicial headnotes or abstracts and a memory containing one or more headnotes requiring classification. The personal computer includes a graphical user interface that concurrently displays one of the headnotes requiring classification, a list of one
- 10 or more candidate classes for the one headnote, at least one classification description associated with one of the listed candidate classes, and at least one classified headnote that is associated with one of the listed candidate classes. The graphical user interface also facilitates user assignment of the one headnote requiring classification to one or more of the listed candidate classes.

- 15 In the exemplary system, the list of candidate classes results from automatically defining and executing a query against the classified headnotes, with the query derived from the one headnote requiring classification. The exemplary system also displays the candidate classes in a ranked order based on measured similarity of corresponding classified headnotes to the headnote
- 20 requiring classification, further assisting the user in assigning the headnote to an appropriate class. Other features of the interface allow the user to reclassify a classified headnote and to define and execute an arbitrary query against the classified headnotes to further assist classification.

Brief Description of Drawings

- 25 Figure 1 is a diagram of an exemplary classification system 100 embodying several aspects of the invention, including a unique graphical user interface 114;
- Figure 2 is a flowchart illustrating an exemplary method embodied in classification system 100 of Figure 1;
- 30 Figure 3 is a diagram illustrating an unclassified document or headnote 300 and a structured query 300' derived from headnote 300 during operation of classification system 100;

Figure 4A is a facsimile of an exemplary graphical user interface 400 that forms a portion of classification system 100.

Figure 4B is a facsimile of exemplary graphical user interface 400 after responding to a user input.

5 Figure 4C is a facsimile of exemplary graphical user interface 400 after responding to another user input.

Figure 5 is a facsimile of an exemplary graphical user interface 500.

Detailed Description of Preferred Embodiments

This description, which references and incorporates the Figures,

10 describes one or more specific embodiments of one or more inventions. These embodiments, offered not to limit but only to exemplify and teach the one or more inventions, are shown and described in sufficient detail to enable those skilled in the art to implement or practice the invention. Thus, where appropriate to avoid obscuring the invention, the description may omit certain information

15 known to those of skill in the art.

The description includes many terms with meanings derived from their usage in the art or from their use within the context of the description. However, as a further aid, the following term definitions are presented.

20 The term "document" refers to any logical collection or arrangement of machine-readable data having a filename.

The term "database" includes any logical collection or arrangement of machine-readable documents.

Figure 1 shows a diagram of an exemplary document classification system 100 for assisting editors in manually classifying electronic documents according to a document classification scheme. The exemplary embodiment assists in the classification of judicial abstracts, or headnotes, according to West Group's Key Number System. For further details on the Key Number System, see West's Analysis of American Law: Guide to the American Digest System, 2000 Edition, West Group, 1999. This text is incorporated herein by reference.

25 30 However, the present invention is not limited to any particular type of documents or type of classification system.

System 100 includes an exemplary personal computer or classification work station 110, an exemplary classified documents database 120, an

exemplary classification system database 130, and an unclassified documents database 140. Though the exemplary embodiment presents work station 110, and databases 120-140 as separate components, some embodiments combine the functionality of these components into a greater or lesser number of components.

5 For example, one embodiment combines databases 120-140 within work station 110, and another embodiment combines database 130 with work station 110 and databases 120 and 140 into a single database.

The most pertinent features of work station 110 include a processing unit 111, a data-storage device 112, a display device 113, a graphical-user interface 10 114, and user-interface devices 115 and 116. In the exemplary embodiment, processor unit 111 includes one or more processors and an operating system which supports graphical-user interfaces. Storage device 112 include one or more electronic, magnetic, and/or optical memory devices. However, other embodiments of the invention, use other types and numbers of processors and 15 data-storage devices. For examples, some embodiment implement one or more portions of system 100 using one or more mainframe computers or servers, such as the Sun Ultra 4000 server. Exemplary display devices include a color monitor and virtual-reality goggles, and exemplary user-interface devices include a keyboard, mouse, joystick, microphone, video camera, body-field sensors, and 20 virtual-reality apparel, such as gloves, headbands, bodysuits, etc. Thus, the invention is not limited to any genus or species of computerized platforms.

Classified documents database 120 includes documents classified according to a classification system. In the exemplary embodiment, database 120 includes an indexed collection of approximately twenty million headnotes 25 spanning the entirety of the West Group's Key Number System. However, some embodiments include an indexed subset of the total collection of classified headnotes. For example, one embodiment indexes headnotes from decisions made within the last 25 years. This reduces the number of headnotes by about half and thus reduces the time necessary to run queries against the the headnotes.

30 Other embodiments further reduce the size of the training collection to include only headnotes specific to the jurisdiction of the query. This is expected not only to result in retrieval of headnotes with greater similarity, but also to further

reduce processing time. Each headnote in the training collection has one or more logically associated Key Number classification codes.

An exemplary indexing procedure entails tokenizing the headnotes, generating transactions, and creating an inverted file. Tokenization entails

5 reading in documents and removing predetermined stop-words, single digits, and stems. The exemplary embodiment uses the Porter stemming algorithm to remove stems. See, M.F. Porter, An Algorithm for Suffix Stripping, Program, 14(3):130-137, July 1980. Single digits are removed since they tend to appear as item markers in enumerations and thus contribute very little to the substance of

10 headnotes.

After tokenization, the procedure generates a transactions for each headnote. A transaction is a tuple grouping a term t , a document identifier n , the frequency of the term t in the document n , and the positions of the term t in document n . Next, the procedure creates an inverted file containing records.

15 The records store the term, the number of documents in the collection that contain the term, and the generated transactions. The inverted file allows efficient access to term information at search time. For further details, see G. Salton, Automatic Text Processing: the Transformation, Analysis and Retrieval of Information by Computer, Addison Wesley, 1989.

20 In addition to an indexed collection of headnotes, database 120 also includes a search engine 121. In the exemplary embodiment, search engine 121 comprises a natural-language search engine, such as the natural language version of WestLaw ® legal search tools. However, other embodiments include other search engines based on the work by H. Turtle, Inference Networks for

25 Document Retrieval, PhD thesis, Computer and Information Science Department, University of Massachusetts, October 1990. Still other embodiments use an Inquiry Retrieval System as described in J.P. Gallan, W.B. Croft, and S.M. Harding, The Inquiry Retrieval System. In Proceedings of the Third International Conference on Database and Expert Systems Applications, 30 pages 78-83, Valencia, Spain, 1992. Springer-Verlag.

Classification system database 130 includes searchable data describing the logical and hierarchical structure of the classification system used in system 100. In the exemplary embodiment, this data describes the approximately

82,000 classes of West Group's Key Number System. Each class description includes its Key Number code, a topic description, and data linking the class to adjacent classes.

Unclassified documents database 140 includes a set of one or more 5 unclassified documents. In the exemplary embodiment, each document is an unclassified headnote or more generally a headnote requiring initial classification or reclassification. Moreover, each headnote has a corresponding judicial opinion. In the exemplary embodiment, the headnotes are determined manually by professional editor. However, other embodiments may determine headnotes 10 automatically using a computerized document summarizer. See for example U.S. Patent 5,708,825 to Bernardo Rafael Sotomayer, which is incorporated herein by reference.

System 100 also includes, within data-storage device 112, classification-aiding software 112a. In the exemplary embodiment, software 112a comprises 15 one or more software modules and operates as a separate application program or as part of the kernel or shell of an operating system. (Software 112a can be installed on work station 110 through a network-download or through a computer-readable medium, such as an optical or magnetic disc, or through other software transfer methods.) In the exemplary embodiment, software 112a 20 enables system 100 to generate graphical-user interface 114 which integrates unclassified headnotes from database 140 with classified headnotes and ranked candidate classes from database 120 and classification system data from database 130 to assist users in manually classifying or reclassifying headnotes.

Figure 2 shows a flow chart 200 of an exemplary classification method at 25 least partly embodied within and facilitated by software 112a. Flow chart 200 includes a number of process blocks 202-214, which are arranged serially in the exemplary embodiment. However, other embodiments of the invention may reorder the blocks, omits one or more blocks, and/or execute two or more blocks in parallel using multiple processors or a single processor organized as two or 30 more virtual machines or subprocessors. Moreover, still other embodiments implement the blocks as one or more specific interconnected hardware or integrated-circuit modules with related control and data signals communicated

between and through the modules. Thus, the exemplary process flow is applicable to software, firmware, and hardware implementations.

The exemplary method begins at process block 202 with automatic or user-directed retrieval of a set of one or more unclassified headnotes from 5 unclassified document database 140. For system embodiments that include two or more classification work stations, a number of sets of unclassified headnotes can be scheduled for classification at particular stations or a set of unclassified headnotes can be queued for sequential distribution to the next available work station. Some embodiments allow the user to define and run a query against the 10 unclassified headnotes and in effect define the set of headnotes he or she will classify or alternatively transfer the set of headnotes to another work station for classification. After retrieval of the unclassified headnotes, execution of the exemplary method then proceeds to block 204.

Block 204 entails defining a query based on one of the headnotes in the 15 set of unclassified headnotes. In the exemplary embodiment, this entails forwarding the one headnote to the natural-language search engine 121 which automatically defines the query using the indexing procedure already applied to index the classified headnotes of database 120. Figure 3 shows the text of a sample headnote 300 and a structured query 300' that search engine 121 derives 20 from it. Although the exemplary embodiment relied on the inherent functionality of its search engine 121 for this query definition some embodiments include a query structuring or definition module within software 112a.

After defining the query, the exemplary method runs, or executes, the 25 query against the classified document database 120, as indicated in block 206. In the exemplary embodiment, search engine 121, which has already defined the query from the unclassified headnote, executes a search based on the query. In executing the search, search engine 121 implements memory-based reasoning, a variant of a *k*-nearest neighbor method. This generally entails retrieving the 30 classified headnotes that are closest to the unclassified headnote, or more precisely the query form of the unclassified headnote, based on some distance function. More particularly, the exemplary embodiment compares the query to each classified headnote in the database, scores all the terms, or concepts, that

each classified headnote has in common with the query, sums the scores of all the common terms, and divides by the total number of query terms in the classified headnote to determine an average score for the classified headnote.

5 In the exemplary embodiment, search engine 121 scores individual terms using the following formula:

$$w(t,d) = 0.4 + 0.6 * tf(t,d) * idf(t),$$

where $w(t,d)$ denotes the weight, or score, for term t in document (or headnote) d ; $idf(t)$ denotes an inverse-document-frequency factor for the term t and $tf(t,d)$ denotes the term-frequency factor for term t in document d . The inverse-
10 document-frequency factor $idf(t)$ is defined as

$$idf(t) = (\log (N) - \log [df(t)]) / \log(N),$$

and the term-frequency factor $tf(t,d)$ for term t in document d is defined as

$$tf(t,d) = 0.5 + 0.5 * \log[f(t,d)] / \log(maxtf),$$

where N is the total number of documents (headnotes) in the collection, $df(t)$ is
15 the number of documents where term t appears, $f(t,d)$ is the number of occurrences of term t in document d , and $maxtf$ is the maximum frequency of any term in document d . The inverse-document-frequency factor (idf) favors (that is, gives greater weight to) terms that are rare in the collection, while the term frequency factor (tf) gives a higher importance to terms that are frequent in
20 the document being scored.

The result of the search is a ranked list of document-score pairs, with each score indicating the similarity between a retrieved classified document and the query. The score is the metric for finding the nearest neighbors. Execution of the method then continues to block 208.

25 Block 208 entails determining the classes associated with a predetermined number k of the top classified headnotes from the ranked list of search results. The k classified headnotes are the k nearest neighbors of the unclassified headnote according to the distance function used in search engine 121. Exemplary values for k include 5, 10, 25, 50, and 100. In the exemplary
30 embodiment, some of the classified headnotes have two or more associated Key Number classes.

After determining all the classes associated with the k classified headnotes most similar to the unclassified headnote, the method executes block

210 which entails transferring the k classified headnotes and their associated class identifiers from classified document database 120 to work station 110.

As block 212 shows, the station 110, or more particular processor unit 111, next determines a ranking for the class identifiers (Key Number classes) 5 associated with the top k classified headnotes. The exemplary embodiment ranks the class identifiers based on their frequencies of occurrence within the set of candidate classes. In other words, each class identifier is ranked based on how many times it appears in the set of candidate classes.

Other embodiments rank the classes based on respective total similarity 10 scores. For a given candidate class, the total similarity score is the sum of the similarity scores for all the headnotes associated with the class. Some embodiments rank the similarity scores for all the headnotes associated with a class, weight the ranks according to a function, and then sum the weighted ranks to determine where to rank the class. Two exemplary rank-weighting functions 15 are:

$$w(r) = 1/r \text{ and}$$

$$w(r) = (1-\varepsilon^*r),$$

where w denotes the weight function and r denotes rank. $\varepsilon = 1/(k+1)$, k being the number of nearest neighbors. Functions such as these give a higher weight to 20 a Key Number class assigned to a document at the top of the retrieved set, and a lower weight when the document is at a lower position.

After ranking the candidate classes, the system executes block 214 which entails displaying on display device 113 (shown in Figure 1) the exemplary graphical user interface 400 which is shown in Figure 4A. Graphical user 25 interface 400 includes concurrently displayed windows or regions 410, 420, 430, 440, and 450.

Window 410 displays the one unclassified headnote, headnote 300 of Figure 3, which was selected or retrieved from classification in block 202 of the exemplary flow chart in Figure 2. Window 420 displays a sorted list or table 422 30 of candidate classes and their corresponding frequencies. A class 422a in list 422 is highlighted in subregion 420a of window 420. Window 430 displays a portion 432a of the classification system hierarchy which includes class 422a. Window 440 displays one or more of the classified headnotes that is similar to

the one unclassified headnote and which has class 422a as one of its assigned classes. Window 450 is an input window for assigning one or more classes to unclassified headnote 412 displayed in window 410.

In operation, interface devices 114-116 of system 100 enable a user to

5 highlight or select one or more of the candidate classes in list 422. For example, a user may point and double click on candidate class 422a (232Ak179) to select the class, or a user may single click on the class to highlight it for further consideration. Selecting, or double-clicking, a class in the list, results in automatic insertion of the class into window 450. The interface not only allows

10 the user to select as many of the classes as desired, but also to manually insert one or more classes, including classes not listed, into window 450. When interface 400 is closed, it prompts the user to save, or in effect, actually assign the one or more classes in window 450 to the headnote in window 410. In response to highlighting class 422a, interface 400 displays subregion 420a of

15 window 420 in reverse-video, that is, by reversing the background and foreground colors of subregion 420a. (Other embodiments use other techniques not only to indicate selection of one of the classes, but also to select one or more of the classes.)

In further response to highlighting a class in list 422 of window 420,

20 classification station 110 (in Figure 1) defines a query based on all or a portion of the highlighted class and runs it against classification system database 130. Database 130 returns one or more classes in the neighborhood of the selected class to station 110, and window 430 displays one or more of these neighborhood classes, as portion 432a, allowing the user to view the highlighted

25 class in context of the classification system, complete with class identifiers and class descriptors.

In addition to responding to highlighting of class 422a by displaying it in context of the classification system in window 430, the interface also displays in window 440 one or more of the classified headnotes that is similar to the

30 headnote being classified. In other words, window 440 displays one of the headnotes, such as headnote 442a, which resulted in the highlighted class 422a being included in list 422. If there are more than one of these headnotes,

window 440 allows the user to view each of them in order from most similar to least similar to the headnote being classified.

Figure 4B shows that the user may also highlight another class, such as class 422b in the list 422 to view this class in context of the classification system 5 in window 430 and to view the classified headnotes associated with the class in window 440. More specifically, window 430 shows a portion 432b of the classification system stored in database 130, and window 440 shows a headnote 442b associated with highlighted class 422b. The interface allows the user to repeat this process with each of the classes in list.

10 Window 430 also includes an enter-query button 434 which the user may invoke to convert window 430 into a query-entry window 430' as shown in Figure 4C. This figure shows an exemplary query 436, which the user has defined to include several terms and/or phrases from or related to unclassified headnote 412 in window 410. The figure also shows that enter-query button 434 15 has been converted to a run-query button 434', which the user may actuate after entering query 436. Actuating the run-query button runs the query against classified documents database 120, and results in representation of interface 400, with an updated list 422' of candidate classes for possible assignment to the unclassified headnote. (Once the user highlights one of the classes in the 20 updated list 422', window 430 will display this class in context of the classification system hierarchy. This user-invocable option of defining and running queries further facilitates classification of headnotes when the candidate classes stemming from the automatically defined queries are unsatisfactory.

When viewing the classified headnotes in window 440, the user may 25 recognize that a particular headnote has been misclassified and thus require reclassification. Thus, window 440 includes a reclassification button 444, which the user can invoke to initiate reclassification of the particular headnote, such as headnote 442b to another class. Invocation of button 444 results in display of window 500 as shown in Figure 5.

30 Window 500 includes a region 510 that displays a headnote 512 that is being reclassified, a region 520 which displays the highlighted class from list 422 that is associated with the headnote, and region 530 displays a ranked list

532 of candidate classes and an input field 534 for entry of new class. Ranked list 532 is developed using the same process used for developing list 422.

Conclusion

In furtherance of the art, the inventors have presented exemplary systems, 5 methods, and software that facilitate the manual classification of documents, particularly judicial headnotes according to a legal classification system, such as West Group's Key Number System. One exemplary system includes a single graphical user interface that concurrently displays one of the headnotes requiring classification, a list of one or more candidate classes for the one headnote, at 10 least one classification description associated with one of the listed candidate classes, and at least one classified headnote that is associated with one of the listed candidate classes. The exemplary interface integrates two or more tools necessary for a user to accurately and efficiently classify judicial headnotes or other documents.

The embodiments described above are intended only to illustrate and teach one or more ways of practicing or implementing the present invention, not to restrict its breadth or scope. The actual scope of the invention, which embraces all ways of practicing or implementing the concepts of the invention, is defined only by the following claims and their equivalents.

Claims

1. A method of classifying one or more documents in a classification scheme including two or more classes, with each class having one or more classified document headnotes, the method comprising:
 - 5 summarizing a particular document to define one or more particular document headnotes;
 - automatically generating a list of one or more of the classes, with each listed class having one or more classified document headnotes which are similar to the particular document headnote; and
 - 10 classifying the particular document or document summary based on the list of classes.
2. A method of classifying one or more documents in a classification scheme including two or more classes, with each class having one or more classified documents, the method comprising:
 - 15 summarizing a particular document to define a particular document summary;
 - automatically generating a list of one or more of the classes, with each listed class having one or more classified documents which are similar to the particular document summary; and
 - 20 classifying the particular document or document summary based on the list of classes.
3. A method of classifying one or more documents in a classification scheme including two or more classes, with each class having one or more classified document summaries, the method comprising:
 - 25 summarizing a particular document to define a particular document summary;
 - automatically generating a list of one or more of the classes, with each listed class having one or more classified document summaries which are similar to the particular document summary; and
 - 30 classifying the particular document based on the list of classes.

4. The method of claim 3, wherein summarizing a particular document comprises manually summarizing the particular document or electronically summarizing the particular document using a computerized text summarizer.
5. The method of claim 3, wherein generating a list of one or more of the classes comprises:
 - defining one or more natural-language or boolean queries based on the particular document summary;
 - performing one or more searches of the classified document summaries based on one or more of the queries, with one or more of the searches yielding one or more found document summaries;
 - ranking the one or more found document summaries based on relative similarity to the particular document summary to define one or more ranked document summaries;
- 10 15 generating the list based on one or more of the ranked document summaries.
6. The method of claim 3, wherein classifying the particular document based on the list of classes comprises manually selecting one or more of the classes using a graphical user interface or automatically selecting one or more of the classes using a predetermined selection procedure.
- 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 105 110 115 120 125 130 135 140 145 150 155 160 165 170 175 180 185 190 195 200 205 210 215 220 225 230 235 240 245 250 255 260 265 270 275 280 285 290 295 300 305 310 315 320 325 330 335 340 345 350 355 360 365 370 375 380 385 390 395 400 405 410 415 420 425 430 435 440 445 450 455 460 465 470 475 480 485 490 495 500 505 510 515 520 525 530 535 540 545 550 555 560 565 570 575 580 585 590 595 600 605 610 615 620 625 630 635 640 645 650 655 660 665 670 675 680 685 690 695 700 705 710 715 720 725 730 735 740 745 750 755 760 765 770 775 780 785 790 795 800 805 810 815 820 825 830 835 840 845 850 855 860 865 870 875 880 885 890 895 900 905 910 915 920 925 930 935 940 945 950 955 960 965 970 975 980 985 990 995 1000 1005 1010 1015 1020 1025 1030 1035 1040 1045 1050 1055 1060 1065 1070 1075 1080 1085 1090 1095 1100 1105 1110 1115 1120 1125 1130 1135 1140 1145 1150 1155 1160 1165 1170 1175 1180 1185 1190 1195 1200 1205 1210 1215 1220 1225 1230 1235 1240 1245 1250 1255 1260 1265 1270 1275 1280 1285 1290 1295 1300 1305 1310 1315 1320 1325 1330 1335 1340 1345 1350 1355 1360 1365 1370 1375 1380 1385 1390 1395 1400 1405 1410 1415 1420 1425 1430 1435 1440 1445 1450 1455 1460 1465 1470 1475 1480 1485 1490 1495 1500 1505 1510 1515 1520 1525 1530 1535 1540 1545 1550 1555 1560 1565 1570 1575 1580 1585 1590 1595 1600 1605 1610 1615 1620 1625 1630 1635 1640 1645 1650 1655 1660 1665 1670 1675 1680 1685 1690 1695 1700 1705 1710 1715 1720 1725 1730 1735 1740 1745 1750 1755 1760 1765 1770 1775 1780 1785 1790 1795 1800 1805 1810 1815 1820 1825 1830 1835 1840 1845 1850 1855 1860 1865 1870 1875 1880 1885 1890 1895 1900 1905 1910 1915 1920 1925 1930 1935 1940 1945 1950 1955 1960 1965 1970 1975 1980 1985 1990 1995 2000 2005 2010 2015 2020 2025 2030 2035 2040 2045 2050 2055 2060 2065 2070 2075 2080 2085 2090 2095 2100 2105 2110 2115 2120 2125 2130 2135 2140 2145 2150 2155 2160 2165 2170 2175 2180 2185 2190 2195 2200 2205 2210 2215 2220 2225 2230 2235 2240 2245 2250 2255 2260 2265 2270 2275 2280 2285 2290 2295 2300 2305 2310 2315 2320 2325 2330 2335 2340 2345 2350 2355 2360 2365 2370 2375 2380 2385 2390 2395 2400 2405 2410 2415 2420 2425 2430 2435 2440 2445 2450 2455 2460 2465 2470 2475 2480 2485 2490 2495 2500 2505 2510 2515 2520 2525 2530 2535 2540 2545 2550 2555 2560 2565 2570 2575 2580 2585 2590 2595 2600 2605 2610 2615 2620 2625 2630 2635 2640 2645 2650 2655 2660 2665 2670 2675 2680 2685 2690 2695 2700 2705 2710 2715 2720 2725 2730 2735 2740 2745 2750 2755 2760 2765 2770 2775 2780 2785 2790 2795 2800 2805 2810 2815 2820 2825 2830 2835 2840 2845 2850 2855 2860 2865 2870 2875 2880 2885 2890 2895 2900 2905 2910 2915 2920 2925 2930 2935 2940 2945 2950 2955 2960 2965 2970 2975 2980 2985 2990 2995 3000 3005 3010 3015 3020 3025 3030 3035 3040 3045 3050 3055 3060 3065 3070 3075 3080 3085 3090 3095 3100 3105 3110 3115 3120 3125 3130 3135 3140 3145 3150 3155 3160 3165 3170 3175 3180 3185 3190 3195 3200 3205 3210 3215 3220 3225 3230 3235 3240 3245 3250 3255 3260 3265 3270 3275 3280 3285 3290 3295 3300 3305 3310 3315 3320 3325 3330 3335 3340 3345 3350 3355 3360 3365 3370 3375 3380 3385 3390 3395 3400 3405 3410 3415 3420 3425 3430 3435 3440 3445 3450 3455 3460 3465 3470 3475 3480 3485 3490 3495 3500 3505 3510 3515 3520 3525 3530 3535 3540 3545 3550 3555 3560 3565 3570 3575 3580 3585 3590 3595 3600 3605 3610 3615 3620 3625 3630 3635 3640 3645 3650 3655 3660 3665 3670 3675 3680 3685 3690 3695 3700 3705 3710 3715 3720 3725 3730 3735 3740 3745 3750 3755 3760 3765 3770 3775 3780 3785 3790 3795 3800 3805 3810 3815 3820 3825 3830 3835 3840 3845 3850 3855 3860 3865 3870 3875 3880 3885 3890 3895 3900 3905 3910 3915 3920 3925 3930 3935 3940 3945 3950 3955 3960 3965 3970 3975 3980 3985 3990 3995 4000 4005 4010 4015 4020 4025 4030 4035 4040 4045 4050 4055 4060 4065 4070 4075 4080 4085 4090 4095 4100 4105 4110 4115 4120 4125 4130 4135 4140 4145 4150 4155 4160 4165 4170 4175 4180 4185 4190 4195 4200 4205 4210 4215 4220 4225 4230 4235 4240 4245 4250 4255 4260 4265 4270 4275 4280 4285 4290 4295 4300 4305 4310 4315 4320 4325 4330 4335 4340 4345 4350 4355 4360 4365 4370 4375 4380 4385 4390 4395 4400 4405 4410 4415 4420 4425 4430 4435 4440 4445 4450 4455 4460 4465 4470 4475 4480 4485 4490 4495 4500 4505 4510 4515 4520 4525 4530 4535 4540 4545 4550 4555 4560 4565 4570 4575 4580 4585 4590 4595 4600 4605 4610 4615 4620 4625 4630 4635 4640 4645 4650 4655 4660 4665 4670 4675 4680 4685 4690 4695 4700 4705 4710 4715 4720 4725 4730 4735 4740 4745 4750 4755 4760 4765 4770 4775 4780 4785 4790 4795 4800 4805 4810 4815 4820 4825 4830 4835 4840 4845 4850 4855 4860 4865 4870 4875 4880 4885 4890 4895 4900 4905 4910 4915 4920 4925 4930 4935 4940 4945 4950 4955 4960 4965 4970 4975 4980 4985 4990 4995 5000 5005 5010 5015 5020 5025 5030 5035 5040 5045 5050 5055 5060 5065 5070 5075 5080 5085 5090 5095 5100 5105 5110 5115 5120 5125 5130 5135 5140 5145 5150 5155 5160 5165 5170 5175 5180 5185 5190 5195 5200 5205 5210 5215 5220 5225 5230 5235 5240 5245 5250 5255 5260 5265 5270 5275 5280 5285 5290 5295 5300 5305 5310 5315 5320 5325 5330 5335 5340 5345 5350 5355 5360 5365 5370 5375 5380 5385 5390 5395 5400 5405 5410 5415 5420 5425 5430 5435 5440 5445 5450 5455 5460 5465 5470 5475 5480 5485 5490 5495 5500 5505 5510 5515 5520 5525 5530 5535 5540 5545 5550 5555 5560 5565 5570 5575 5580 5585 5590 5595 5600 5605 5610 5615 5620 5625 5630 5635 5640 5645 5650 5655 5660 5665 5670 5675 5680 5685 5690 5695 5700 5705 5710 5715 5720 5725 5730 5735 5740 5745 5750 5755 5760 5765 5770 5775 5780 5785 5790 5795 5800 5805 5810 5815 5820 5825 5830 5835 5840 5845 5850 5855 5860 5865 5870 5875 5880 5885 5890 5895 5900 5905 5910 5915 5920 5925 5930 5935 5940 5945 5950 5955 5960 5965 5970 5975 5980 5985 5990 5995 6000 6005 6010 6015 6020 6025 6030 6035 6040 6045 6050 6055 6060 6065 6070 6075 6080 6085 6090 6095 6100 6105 6110 6115 6120 6125 6130 6135 6140 6145 6150 6155 6160 6165 6170 6175 6180 6185 6190 6195 6200 6205 6210 6215 6220 6225 6230 6235 6240 6245 6250 6255 6260 6265 6270 6275 6280 6285 6290 6295 6300 6305 6310 6315 6320 6325 6330 6335 6340 6345 6350 6355 6360 6365 6370 6375 6380 6385 6390 6395 6400 6405 6410 6415 6420 6425 6430 6435 6440 6445 6450 6455 6460 6465 6470 6475 6480 6485 6490 6495 6500 6505 6510 6515 6520 6525 6530 6535 6540 6545 6550 6555 6560 6565 6570 6575 6580 6585 6590 6595 6600 6605 6610 6615 6620 6625 6630 6635 6640 6645 6650 6655 6660 6665 6670 6675 6680 6685 6690 6695 6700 6705 6710 6715 6720 6725 6730 6735 6740 6745 6750 6755 6760 6765 6770 6775 6780 6785 6790 6795 6800 6805 6810 6815 6820 6825 6830 6835 6840 6845 6850 6855 6860 6865 6870 6875 6880 6885 6890 6895 6900 6905 6910 6915 6920 6925 6930 6935 6940 6945 6950 6955 6960 6965 6970 6975 6980 6985 6990 6995 7000 7005 7010 7015 7020 7025 7030 7035 7040 7045 7050 7055 7060 7065 7070 7075 7080 7085 7090 7095 7100 7105 7110 7115 7120 7125 7130 7135 7140 7145 7150 7155 7160 7165 7170 7175 7180 7185 7190 7195 7200 7205 7210 7215 7220 7225 7230 7235 7240 7245 7250 7255 7260 7265 7270 7275 7280 7285 7290 7295 7300 7305 7310 7315 7320 7325 7330 7335 7340 7345 7350 7355 7360 7365 7370 7375 7380 7385 7390 7395 7400 7405 7410 7415 7420 7425 7430 7435 7440 7445 7450 7455 7460 7465 7470 7475 7480 7485 7490 7495 7500 7505 7510 7515 7520 7525 7530 7535 7540 7545 7550 7555 7560 7565 7570 7575 7580 7585 7590 7595 7600 7605 7610 7615 7620 7625 7630 7635 7640 7645 7650 7655 7660 7665 7670 7675 7680 7685 7690 7695 7700 7705 7710 7715 7720 7725 7730 7735 7740 7745 7750 7755 7760 7765 7770 7775 7780 7785 7790 7795 7800 7805 7810 7815 7820 7825 7830 7835 7840 7845 7850 7855 7860 7865 7870 7875 7880 7885 7890 7895 7900 7905 7910 7915 7920 7925 7930 7935 7940 7945 7950 7955 7960 7965 7970 7975 7980 7985 7990 7995 8000 8005 8010 8015 8020 8025 8030 8035 8040 8045 8050 8055 8060 8065 8070 8075 8080 8085 8090 8095 8100 8105 8110 8115 8120 8125 8130 8135 8140 8145 8150 8155 8160 8165 8170 8175 8180 8185 8190 8195 8200 8205 8210 8215 8220 8225 8230 8235 8240 8245 8250 8255 8260 8265 8270 8275 8280 8285 8290 8295 8300 8305 8310 8315 8320 8325 8330 8335 8340 8345 8350 8355 8360 8365 8370 8375 8380 8385 8390 8395 8400 8405 8410 8415 8420 8425 8430 8435 8440 8445 8450 8455 8460 8465 8470 8475 8480 8485 8490 8495 8500 8505 8510 8515 8520 8525 8530 8535 8540 8545 8550 8555 8560 8565 8570 8575 8580 8585 8590 8595 8600 8605 8610 8615 8620 8625 8630 8635 8640 8645 8650 8655 8660 8665 8670 8675 8680 8685 8690 8695 8700 8705 8710 8715 8720 8725 8730 8735 8740 8745 8750 8755 8760 8765 8770 8775 8780 8785 8790 8795 8800 8805 8810 8815 8820 8825 8830 8835 8840 8845 8850 8855 8860 8865 8870 8875 8880 8885 8890 8895 8900 8905 8910 8915 8920 8925 8930 8935 8940 8945 8950 8955 8960 8965 8970 8975 8980 8985 8990 8995 9000 9005 9010 9015 9020 9025 9030 9035 9040 9045 9050 9055 9060 9065 9070 9075 9080 9085 9090 9095 9100 9105 9110 9115 9120 9125 9130 9135 9140 9145 9150 9155 9160 9165 9170 9175 9180 9185 9190 9195 9200 9205 9210 9215 9220 9225 9230 9235 9240 9245 9250 9255 9260 9265 9270 9275 9280 9285 9290 9295 9300 9305 9310 9315 9320 9325 9330 9335 9340 9345 9350 9355 9360 9365 9370 9375 9380 9385 9390 9395 9400 9405 9410 9415 9420 9425 9430 9435 9440 9445 9450 9455 9460 9465 9470 9475 9480 9485 9490 9495 9500 9505 9510 9515 9520 9525 9530 9535 9540 9545 9550 9555 9560 9565 9570 9575 9580 9585 9590 9595 9600 9605 9610 9615 9620 9625 9630 9635 9640 9645 9650 9655 9660 9665 9670 9675 9680 9685 9690 9695 9700 9705 9710 9715 9720 9725 9730 9735 9740 9745 9750 9755 9760 9765 9770 9775 9780 9785 9790 9795 9800 9805 9810 9815 9820 9825 9830 9835 9840 9845 9850 9855 9860 9865 9870 9875 9880 9885 9890 9895 9900 9905 9910 9915 9920 9925 9930 9935 9940 9945 9950 9955 9960 9965 9970 9975 9980 9985 9990 9995 10000 10005 10010 10015 10020 10025 10030 10035 10040 10045 10050 10055 10060 10065 10070 10075 10080 10085 10090 10095 10100 10105 10110 10115 10120 10125 10130 10135 10140 10145 10150 10155 10160 10165 10170 10175 10180 10185 10190 10195 10200 10205 10210 10215 10220 10225 10230 10235 10240 10245 10250 10255 10260 10265 10270 10275 10280 10285 10290 10295 10300 10305 10310 10315 10320 10325 10330 10335 10340 10345 10350 10355 10360 10365 10370 10375 10380 10385 10390 10395 10400 10405 10410 10415 10420 10425 10430 10435 10440 10445 10450 10455 10460 10465 10470 10475 10480 10485 10490 10495 10500 10505 10510 10515 10520 10525 10530 10535 10540 10545 10550 10555 10560 10565 10570 10575 10580 10585 10590 10595 10600 10605 10610 10615 10620 10625 10630 10635 10640 10645 10650 10655 10660 10665 10670 10675 10680 10685 10690 10695 10700 10705 10710 10715 10720 10725 10730 10735 10740 10745 10750 10755 10760 10765 10770 10775 10780 10785 10790 10795 10800 10805 10810 10815 10820 10825 10830 10835 10840 10845 10850 10855 10860 10865 10870 10875 10880 10885 10890 10895 10900 10905 10910 10915 10920 10925 10930 10935 10940 10945 10950 10955 10960 10965 10970 10975 10980 10985 10990 10995 11000 11005 11010 11015 11020

8. A method of classifying one or more documents, comprising providing a classification scheme including two or more classes, with each class having one or more classified document summaries logically associated with it;
 - 5 summarizing a particular document to define a particular document summary; automatically generating a list of one or more of the classes, with each listed class having one or more classified document summaries which are similar to the particular document summary; and
 - 10 classifying the particular document based on the list of classes.
9. The method of claim 8, wherein summarizing a particular document comprises manually summarizing the particular document or electronically summarizing the particular document using a computerized text summarizer.
- 15
10. The method of claim 8, wherein generating a list of one or more of the classes comprises:
 - defining one or more natural-language or boolean queries based on the particular document summary;
 - 20 performing one or more searches of the classified document summaries based on one or more of the queries, with one or more of the searches yielding one or more found document summaries;
 - ranking the one or more found document summaries based on relative similarity to the particular document summary to define one or
 - 25 more ranked document summaries;
 - generating the list based on one or more of the ranked document summaries.
11. The method of claim 8, wherein classifying the particular document
- 30 based on the list of classes comprises manually selecting one or more of the classes using a graphical user interface or automatically selecting one or more of the classes using a predetermined selection procedure.

12. The method of claim 8, further comprising adding one or more classes to the classification scheme, with each added class having one or more classified document summaries logically associated with it.
- 5 13. The method of claim 8, wherein each class has an associated legal concept and the particular document is a judicial opinion or secondary legal source.
- 10 14. The method of claim 8, wherein the classification scheme conforms at least in part with a version of the West Key Numbering System.
- 15 15. A computer-readable magnetic, electronic, or optical medium comprising computer-executable instructions for:
 - causing a computer to read at least part of a classification scheme into memory, the classification scheme including two or more classes, with each class having one or more classified document summaries logically associated with it;
 - causing the computer to summarize in memory a particular document to define a particular document summary;
 - 20 causing the computer to generate a list in memory of one or more of the classes, with each listed class having associated with it one or more classified document summaries which are similar to the particular document summary; and
 - causing the computer to classify the particular document based on the list of classes.
- 25 16. The medium of claim 15, wherein the instructions for summarizing a particular document comprises instructions for causing the computer to weigh the lexical content of the document.
- 30 17. The medium of claim 15, wherein the instructions for generating a list of one or more of the classes comprises instructions for:

causing the computer to define one or more natural-language or boolean queries based on the particular document summary;

causing the computer to perform one or more searches of the classified document summaries based on one or more of the queries, with

5 one or more of the searches yielding one or more found document summaries;

causing the computer to rank the one or more found document summaries based on relative similarity to the particular document summary to define one or more ranked document summaries; and

10 causing the computer to generate the list based on one or more of the ranked document summaries.

18. The medium of claim 15, wherein the instructions for classifying the particular document based on the list of classes comprises instructions for

15 causing the computer to facilitate manual selection one or more of the classes using a graphical user interface or instructions for causing the computer to automatically select one or more of the classes using a predetermined selection procedure.

20 19. The medium of claim 15, further comprising instructions for manually or automatically adding one or more classes to the classification scheme, with each added class having one or more classified document summaries logically associated with it.

25 20. The medium of claim 15, wherein each class has an associated legal concept and the particular document is a judicial opinion.

21. The medium of claim 15, wherein the classification scheme conforms at least in part with a version of the West Key Numbering System.

30

22. A system for classifying one or more documents in a classification scheme including two or more classes, with each class having one or more classified document summaries, the system comprising:

means for summarizing a particular document to define a particular document summary;

means for automatically generating a list of one or more of the classes, with each listed class having one or more classified document summaries which are similar to the particular document summary;

5 and

means for classifying the particular document based on the list of classes.

23. The system of claim 22, wherein the means for summarizing, the means for automatically generating a list, and the means for classifying exist as software module in a memory coupled to one or more computer processors or within various parts of a mainframe computer or within a SUN Ultra 4000 Server.

10 24. The system of claim 22, wherein the means for summarizing comprises the summarizer described in United States Patent 5,708,825 to Bernardo Rafael Sotomayer, which is incorporated herein by reference.

15 25. A system for classifying one or more documents, comprising means for providing a classification scheme including two or more classes, with each class having one or more classified document summaries logically associated with it;

means for summarizing a particular document to define a particular document summary;

20 25 means for automatically generating a list of one or more of the classes, with each listed class having one or more classified document summaries which are similar to the particular document summary;

and

means for classifying the particular document based on the list of classes.

30 26. A graphical user interface for aiding manual classification of one or more documents in a document classification system having two or more classes, the interface comprising:

means for displaying at least a portion of one of the documents; and
means for displaying information identifying one or more of the classes
as candidate classes.

5 27. The graphical user of claim 26, wherein each document is a headnote, the
headnote associated with a judicial opinion.

10 28. A graphical user interface for aiding manual classification of one or more
documents in a document classification system having two or more classes, the
interface comprising:

means for displaying at least a portion of one of the documents;
means for displaying information identifying one or more of the classes
as candidate classes; and

15 means for displaying a logical relationship between at least one of the
candidate classes and another class in the document classification
system.

20 29. A graphical user interface for aiding manual classification of documents
according to a document classification system having two or more classes, the
interface comprising:

means for displaying at least a portion of one of the documents;
means for displaying information identifying one or more of the classes
as candidate classes for the one of the documents;

25 means for displaying a logical relationship between at least one of the
candidate classes and another class in the document classification
system; and

means for displaying at least one classified document associated with one
of the candidate classes.

30

30. A method for aiding manual classification of documents according to a
document classification system having two or more classes, the method
comprising:

displaying at least a portion of one of the documents;
displaying information identifying one or more of the classes as
candidate classes for the one of the documents, the information
displayed concurrently with the portion of the one or more
5 documents;
displaying a logical relationship between at least one of the candidate
classes and another class in the document classification system,
the logical relationship displayed concurrent with the the
information; and
10 displaying at least a portion of one classified document associated with
one of the candidate classes, the portion of the one classified
documents displayed concurrent with the logical relationship.

31. The method of claim 30, wherein the logical relationship is a
15 hierarchical relationship of at least one the candidate classes to one or more
adjacent classes in the document classification system.

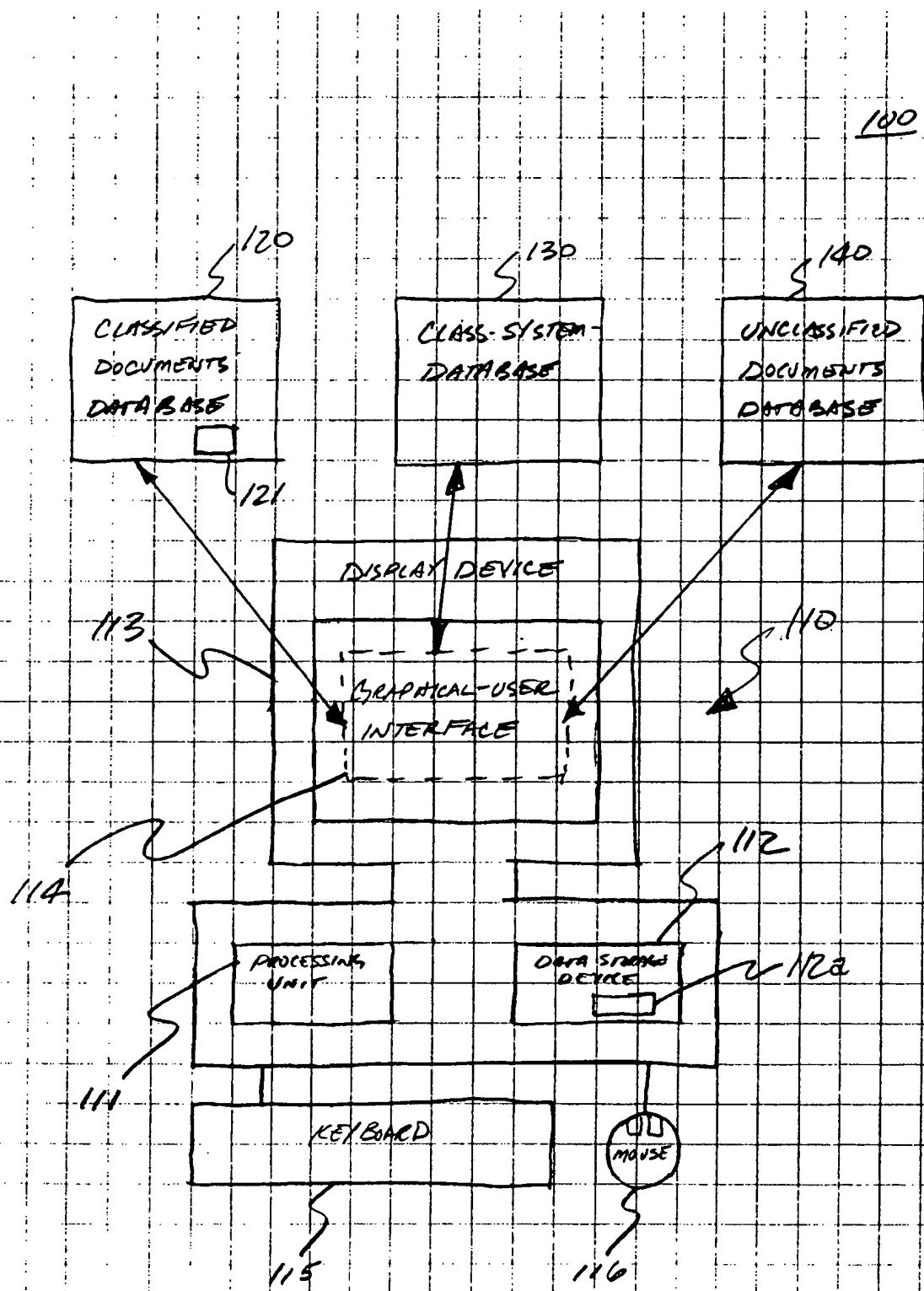
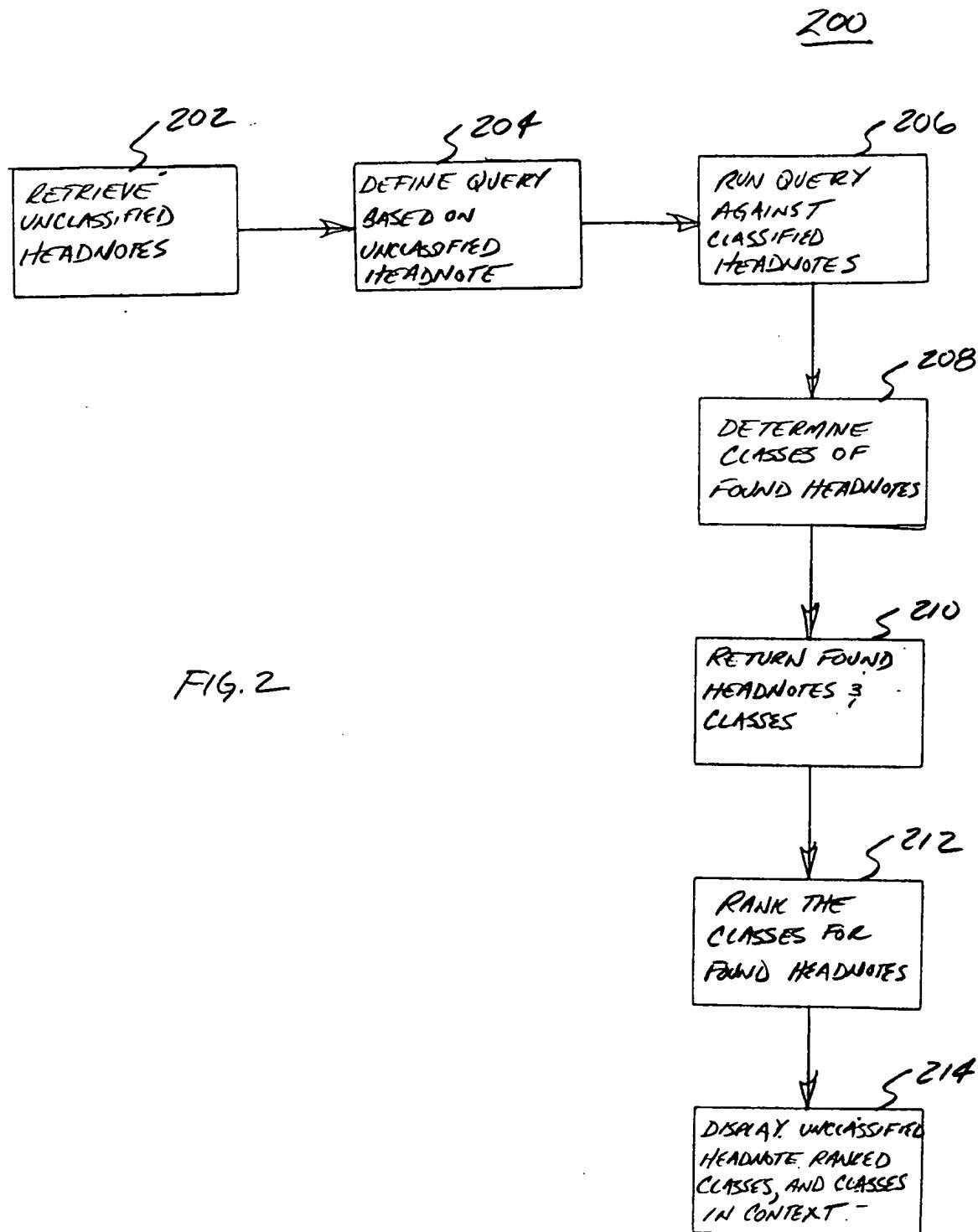


FIG. 1



Employer's duty, under NLRA, to bargain in good faith includes obligation to provide union with information relevant to collective bargaining process in certain circumstances.

300

300'

```
sum(  
    employer  
    duty  
    nlra  
    bargain  
    #phrase(  
        good  
        faith  
    )  
    includes  
    obligation  
    provide  
    union  
    information  
    relevant  
    #phrase(  
        collective  
        bargaining  
    )  
    process  
    certain  
    circumstances  
    #phrase(  
        national  
        labor  
        relations  
        act  
    )  
    amended  
)
```

FIG. 3

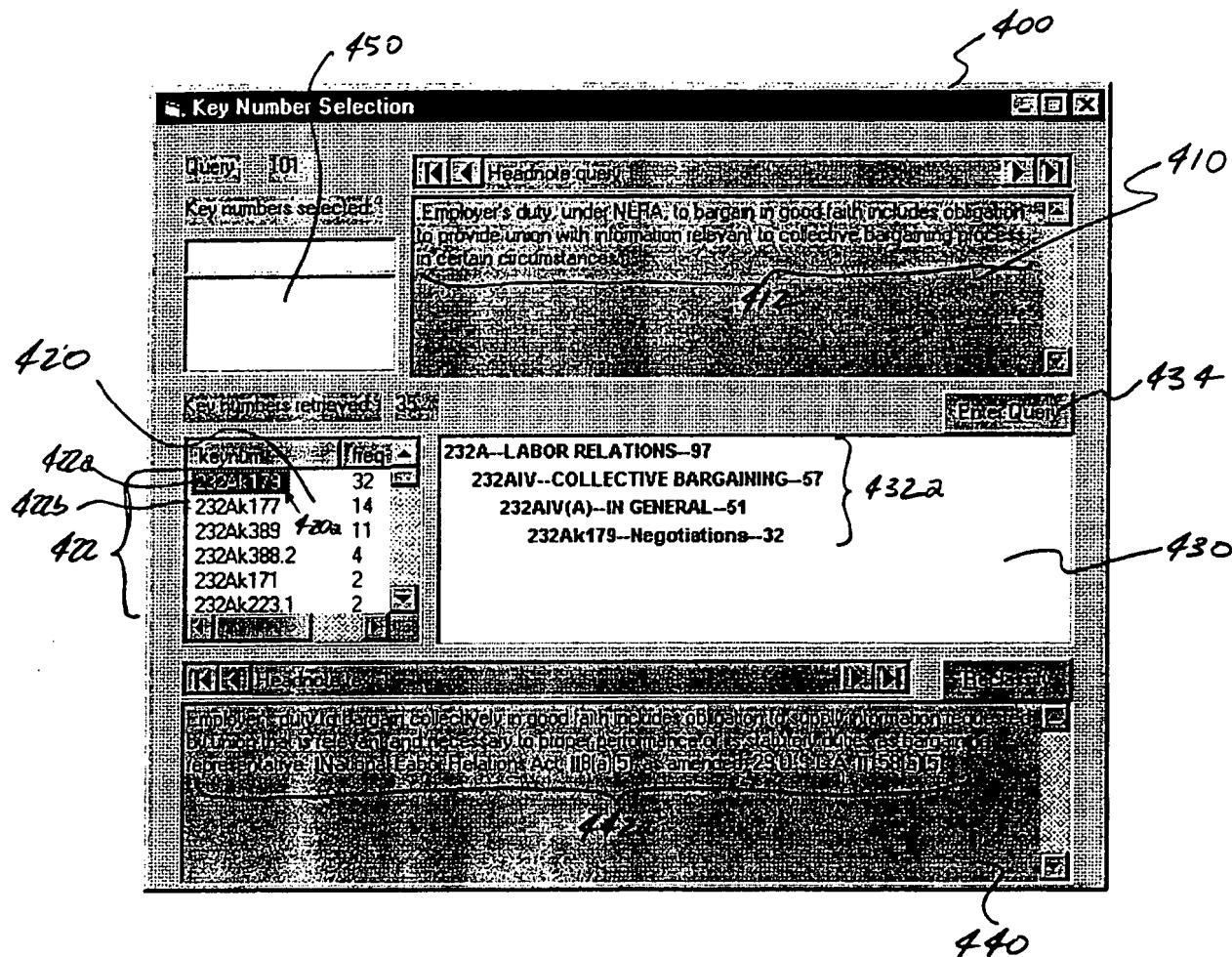


FIG. 4A

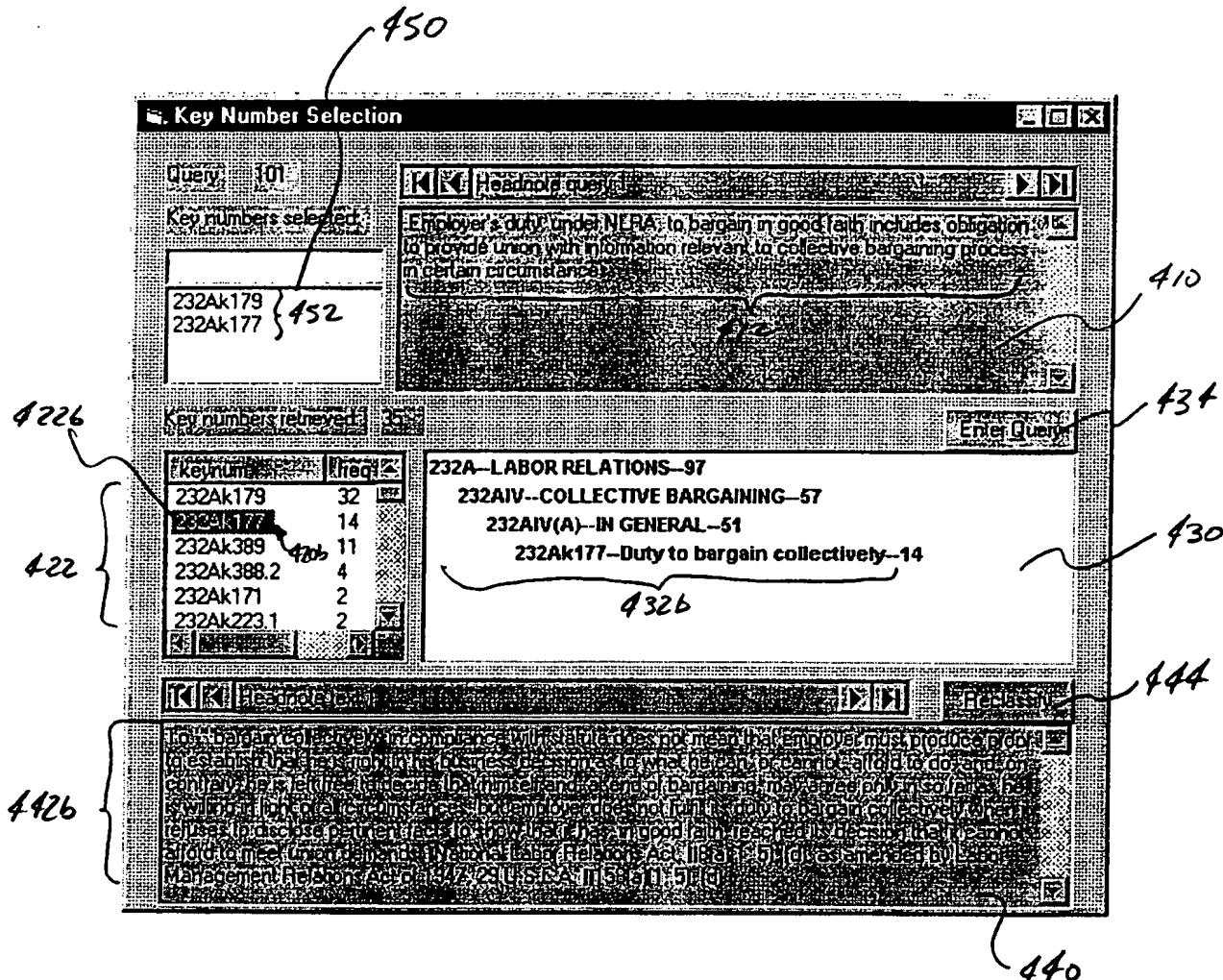


FIG. 4B

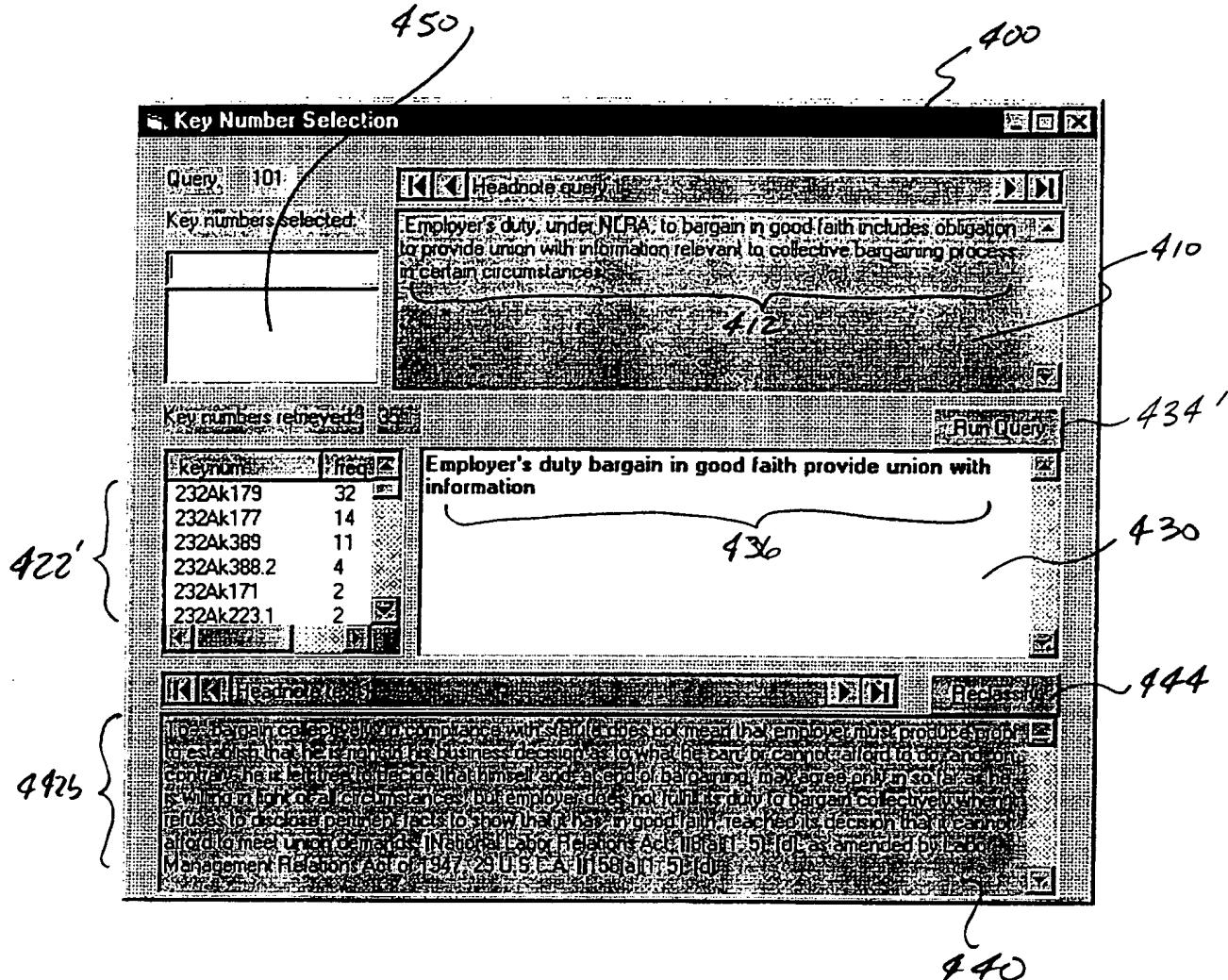


FIG. 4C

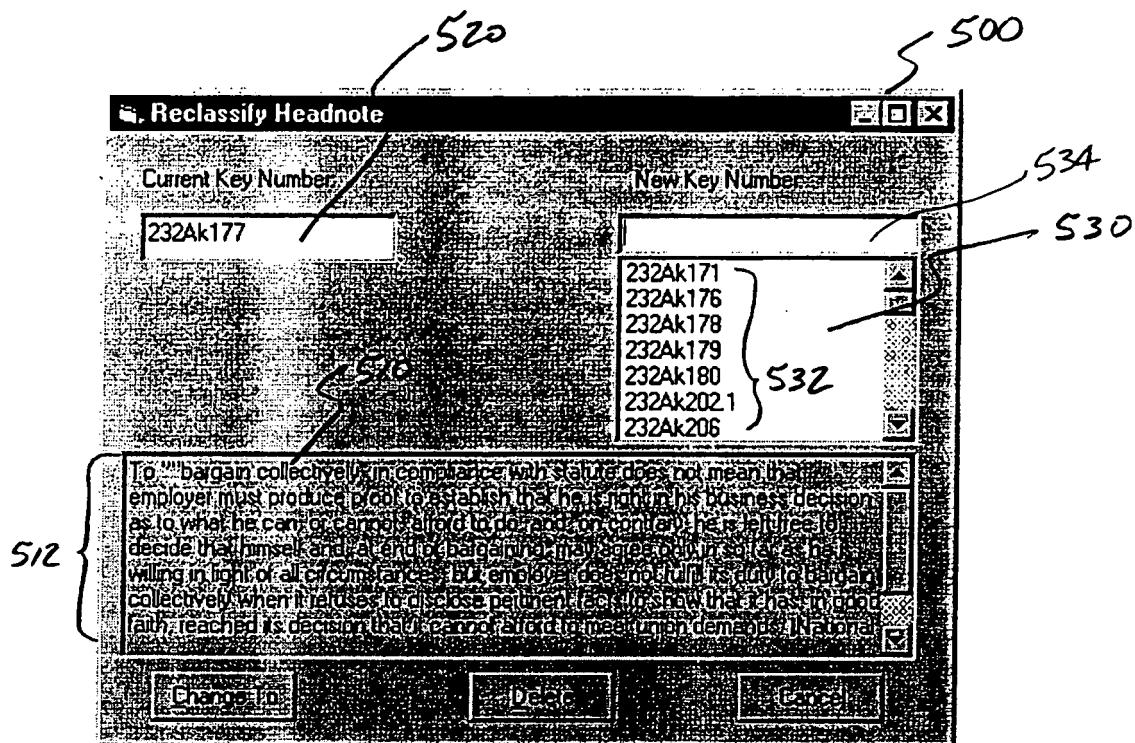


FIG. 5

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US00/12386

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : G06F 17/30

US CL : 707/2, 3, 5, 514, 523, 529, 907; 364/300, 400, 401, 409.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 707/2, 3, 5, 514, 523, 529, 907; 364/300, 400, 401, 409.

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WEST

Search terms: document classification, headnotes or abstracts, summarize, summary list, ranking, query, search.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|-----------|---|-----------------------|
| Y | US 5,794,236 A (MEHRLE) 11 August 1998, col. 2, lines 10-27 and lines 41-67, col. 3, lines 1-67, col. 4, lines 1-23, col. 5, lines 7-24 and lines 49-67, col. 6, lines 1-64, col. 7, lines 29-44, col. 8, lines 43-53, col. 9, lines 45-67, col. 10, lines 1-33. | 1-31 |
| Y | US 5,815,392 A (BENNETT et al) 29 September 1998, col. 4, lines 36-57, col. 5, lines 26-67, col. 6, lines 1-21, col. 7, lines 10-23 and lines 55-67, col. 8, lines 1-34, col. 9, lines 6-67, col. 9, lines 1-12 and lines 34-45, col. 10, lines 22-36, col. 12, lines 7-26, col. 14, lines 20-32, col. 15, lines 37-54, col. 16, lines 24-56, col. 17, lines 50-67, col. 18, lines 1-37, and col. 19, lines 5-16 and lines 55-61. | 1-31 |

Further documents are listed in the continuation of Box C. See patent family annex.

| | | |
|---|-----|--|
| * Special categories of cited documents: | *T* | later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention |
| *A* document defining the general state of the art which is not considered to be of particular relevance | *X* | document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone |
| *B* earlier document published on or after the international filing date | *Y* | document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art |
| *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) | *&* | document member of the same patent family |
| *O* document referring to an oral disclosure, use, exhibition or other means | | |
| *P* document published prior to the international filing date but later than the priority date claimed | | |

| | |
|---|--|
| Date of the actual completion of the international search | Date of mailing of the international search report |
| 07 JULY 2000 | 02 AUG 2000 |

| | |
|---|--|
| Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230 | Authorized officer KIM VU Telephone No. (703) 305-3393 |
|---|--|

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US00/12386

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|-----------|---|-----------------------|
| Y | US 5,708,825 A (SOTOMAYOR) 13 January 1998, col. 4, lines 34-43, col. 5, lines 43-67, col. 6, lines 1-17, col. 7, lines 11-47, col. 8, lines 26-43, col. 11, lines 19-59, col. 12, lines 59-67, col. 13, lines 1-27 and lines 47-67, col. 14, lines 1-67, col. 15, lines 1-19 and lines 48-67, col. 16, lines 1-12 and lines 25-61, col. 17, lines 35-67, col. 18, lines 1-36, col. 32, lines 8-15, and col. 34, lines 37-49. | 1-31 |